

Accessing Accumulated Knowledge in Online Community Question Answering Services

Jun Choi Lee

Universiti Malaysia Sarawak, Kota Samarahan, Malaysia
jclee@fit.unimas.my

Yu-N Cheah

Universiti Sains Malaysia, Penang, Malaysia
yncheah@usm.my

Abstract:

Online Community Question Answering Service (CQAs) provides an online platform for people to seek answers from the public. It is a problem solving method using public intelligent, which also known as crowd sourcing. Each response presented in CQAs is generated by public. Most of this user generated responses are aimed to answer the question posted in the CQAs. This means the repository for CQAs is a huge knowledge pool on the web. This knowledge can serves as knowledge source for intelligent application such as Question Answering. There are two major problems in harvesting this knowledge: first problem is gaining access to the question and answers in the CQAs, second problem is identifying question and answers that content crucial or useful information for other knowledge process. This paper presents an Information Retrieval framework to query and access to the knowledge accumulated in the CQAs. This framework includes a method to query a selected CQAs and a method to filter out quality content from the retrieved data using semantic relatedness. The framework is showcased through an automatic Question Answering application. The application obtained candidate answers from selected CQAs using the proposed framework before determine the final answers through answer ranking method.

Keywords: *Online Community Question Answering Services; Information Retrieval Framework; Question Answering Model*

Introduction

Online Community QA Services (CQAs) is an online service for user to seek answers. Different to traditional question and answering applications, these answers are generated by other user. The person who asked the question can then select from the various answers provided by other user as the best answer for answering their question. This kind of services allows users to share their knowledge or experience of solving certain real-life problem. Thus the CQAs Services holds a lot of unstructured knowledge that is useful to other knowledge processes such as Question Answering (QA), opinion mining, and decision support system.

Obtaining the knowledge from these CQAs faces several challenges. The first challenge is to query and retrieve information from a particular CQAs Services. The live data storage for